DEFENCE
SCIENCE & TECHNOLOGY

Non-linear and Robust Filtering: From the Kalman Filter to the Particle Filter

Jason Ford

DSTO-TR-1301

Non-linear and Robust Filtering: From the Kalman Filter to the Particle Filter

Jason J. Ford

Weapons Systems Division
Aeronautical and Maritime Research Laboratory

DSTO–TR–1301

ABSTRACT

This report presents a review of recent non-linear and robust filtering results for stochastic systems. We focus on stability and robustness issues that arise in the filtering of real systems. Issues such as numeric stability and the effect of non-linearity are also considered.

The report begins by introducing the famous Kalman filtering problem before proceeding to introduce the extended Kalman filter and related stability results. Robust forms of the Kalman filter and extended Kalman filter are also considered and finally a particle filtering approach is presented.

The report is intended to lead readers with a familiarity of the Kalman filtering problem through some of the more important recent (and not so recent) results on stability and robust filters in non-linear filtering problems.

APPROVED FOR PUBLIC RELEASE

Published by

DSTO Aeronautical and Maritime Research Laboratory
506 Lorimer St,
Fishermans Bend, Victoria, Australia 3207

Telephone: (03) 9626 7000
Facsimile: (03) 9626 7999

AR No. AR 012-300
April, 2002

APPROVED FOR PUBLIC RELEASE

Non-linear and Robust Filtering: From the Kalman Filter to the Particle Filter

EXECUTIVE SUMMARY

In the context of signal analysis, filtering is the process of separating a signal of interest from other signals, termed noise signals. All filtering of signals is based on assumptions regarding the nature of the interesting signal and the noise signals. Often filters are used even when it is known that these assumptions do not hold. In this report we present a review of standard filtering techniques, examine techniques for improving the numeric stability of filters, present stochastic stability results and examine filters that mitigate for the possibility of modelling errors.

Modern guided weapons are becoming required to operate in highly non-linear and time-varying situations. In these types of engagements, assumptions of linearity no longer hold and model uncertainties in the form of unmeasured aerodynamic coefficients and complex non-linear aerodynamics are common. Filtering is an important sub-system of any guidance loop that is needed to estimate required engagement information. For these reasons, a full understanding of the effects of non-linearity and model uncertainty on filtering solutions is required.

This report provides a review of existing stochastic filtering results from the well known Kalman filter and non-linear and robust generalisation of the Kalman filter through to new particle filter approaches for non-linear state estimation problems. Several of these filters are tested in simulations of a typical interceptor-target engagement to examine them in a guided weapon context.

An improved understanding of filtering techniques is required to aid support of present upgrade programs involving the guidance loops of new standoff missile systems. This understanding is necessary for the support of future weapon procurement and upgrade programs.

Author

Jason J. Ford
Weapons Systems Division

Jason Ford joined the Guidance and Control group in the Weapons Systems Division in February 1998. He received B.Sc. (majoring in Mathematics) and B.E. degrees from the Australian National University in 1995. He also holds the PhD (1998) degree from the Australian National University. His thesis presents several new on-line locally and globally convergent parameter estimation algorithms for hidden Markov models (HMMs) and linear systems.

His research interests include: adaptive signal processing, adaptive parameter estimation, risk-sensitive techniques, non-linear filtering and applications of HMM and Kalman filtering techniques.

Contents

1 Introduction

The purpose of filtering is to separate one thing from another. In the context of signal analysis, filtering is the process of separating a signal of interest from other signals termed, noise signals. All filtering of signals is based on assumptions regarding the nature of the interesting signal and the noise signals. A system model is a collection of assumptions that describe the relationship between the signal and noise. Ideally the assumed system model should perfectly describe the true system but this is not often the case. Often filters are used even when it is known that these assumptions do not hold. In this report we examine how the performance of a filtering solution is effected by mismatch between the assumed model and the real system.

The most famous and commonly used assumptions are that the system is linear Gauss-Markov and that the noises are Gaussian. It has been shown that under these assumptions the Kalman filter is the optimal filter for separating the interesting signals generated by a linear Gauss-Markov model observed in Gaussian noise [1]. Because the Kalman filter is an optimal filter (in a minimum mean squares sense) and is a finite dimensional solution (can be implemented using a finite number of statistics which can be calculated using a finite number of recursions) it has been applied to a large variety of filtering problems. The continuing success of the Kalman filter in many applications has encouraged its use even in situations where it is clear that the system is non-linear.

An important implementation consideration is the numeric stability of any filtering algorithm. Although the Kalman filter is optimal in a minimum mean least squares sense, it has been shown that it can be numerically unstable when implemented on a finite word length processor [2]. Such numeric issues are probably not as significant today as they where in the early 1960s because of the advances on computer technology; however, in some situations the numeric problems of the standard form of the Kalman filter may still be an issue. For example, in very low noise situations the numeric errors due to finite-precision arithmetic can result in the standard form of the Kalman filter becoming unstable [2]. This report gives details of how to avoid some of the numeric problems inherent in the standard form of the Kalman filter by implementing a numerically stable form such as the U-D covariance form of the Kalman filter.

Apart from the Kalman filter, there are very few finite-dimensional optimal filters for stochastic filtering problems. For general non-linear problems, when a finite-dimensional optimal filter is not possible, sub-optimal numeric or approximate approaches must be used. The simplest approach is to linearise the non-linear model about various operating points and to use an extension of the Kalman filter, known as the extended Kalman filter (EKF). Details of the extended Kalman filter are given in this report as well as some recent stability results that establish under what conditions the extended Kalman filter will give reasonable filtering solutions.

Both the Kalman filter and the extended Kalman filter assume that true system model is known with certainty for design purposes. This is unlikely in most practical situations. It is more likely that the system model will not be known with complete certainty. Kalman filtering theory does not establish how the Kalman filter performs if system assumptions are incorrect. In fact, even a very slight error in the system model can result in poor filter performance. The question of how to design filters when there is some uncertainty in the system model is addressed by robust filtering theory.

In this report we present some recent results describing the robust linear Kalman filter. The robust Kalman filter (or set-valued state estimator) allows state estimation for any member of a structured set of model (with bounded rather than Gaussian noise terms). The filter is termed a set-valued state estimator because it determines the set of possible values that can taken on by the state process given the true system is one from a structured set of models. That is, it determines every possible state sequence that can match the observations. Under the model uncertainty, any of these sequences are possible and the set of estimated state sequence is usually represented by a sequence in the middle of the set.

Although not considered in this report, there are several other approaches that allow for some model uncertainty. Risk-sensitive filtering is an optimal filtering approach in which the filtering criteria is modified in a particular way to mitigate the possibility of modelling error. Risk-sensitive filters tend to have better conditional mean error performance than H^∞ filtering (but less than optimal conditional mean filters) and have some capability to handle modelling errors [11, 12, 13]. An alternative approach is mixed criteria state estimation which is examined for hidden Markov models (HMM) in [14].

The Kalman filter, the extended Kalman filter and the robust filter approaches all estimate state information such as mean and variance. An alternative approach is to consider the evolution of the probability density functions (PDF) directly. The integrals describing the evolution of these probability density functions generally do not have analytic solutions (except in simple cases corresponding to the Kalman filter) and solutions must be obtained through numeric methods. One interesting numeric approach that has recently appeared is the particle filter approach.

In the particle filter approach the probability density function of the system state is approximated by a system of discrete points (or particles) which evolve according to the system equations [16]. At any time point, the set of particles can be used to approximate the probability density function of the state. It can be shown that as the number of particle increases to infinity, the approximation approaches the true probability density function [16]. These particle filter approaches have the advantage that they work for any general non-linear system description. Closely related to the particle filter approach are importance sampling approaches [26] and Monte-Carlo simulation techniques [27]. In this report we present the particle filter approach but no details of these related techniques are provided.

Another technique for estimating the PDF directly is via an HMM approximation. If the state space is naturally bounded then it is possible to approximate the state space by finite regions and use the powerful signal processing tool developed for hidden Markov models to estimate the PDF. An outline of an HMM approximation technique is given.

This report is organised as follows: In Section 2 the Kalman filtering problem is presented and the optimal solution is given. In Section 3 a numerically stable form of the Kalman filter known as the U-D covariance factorisation of the Kalman filter is presented. The extended Kalman filter along with stability results are then presented in Section 4. In Section 5 some robust filtering results are presented. The particle filtering approach is represented in Section 6. Some details of hidden Markov model approximations are then given. Some application examples are presented in Section 7 which compare the various algorithms presented in this report. Finally, in Section 8 some conclusions are presented.

2 Discrete-Time Kalman Filter Problem

This section introduces a model for a discrete-time linear Gauss-Markov system and the optimal filter for the state variable. This optimal filter is known as the Kalman filter.

Consider a stochastic process x_k taking values in R^N. Assume that the dynamics of x_k are described for $k \in Z^+$, where Z^+ is the set of non-negative integers, by

$$x_{k+1} = A_k x_k + B_k w_k \tag{2.1}$$

where $A_k \in R^{(N \times N)}$ is called the state transition matrix, $B_k \in R^{(N \times S)}$ and $w_k \in R^{(S \times 1)}$ is a sequence of independent and identically distributed (*iid*) $N(0, Q_k)$ vector random variables called the process noise. The matrix Q_k is a non-negative definite symmetric matrix. We assume that x_0 is a $N(\hat{x}_0, P_0)$ vector random variable. For simplicity we assume that there is no input into this system but the results in this report can easily be extended to a system with a measured input signal.

Further suppose that x_k is observed indirectly via the vector measurement process $y_k \in R^{(M \times 1)}$ described as follows:

$$y_k = C_k x_k + v_k \tag{2.2}$$

where $C_k \in R^{(M \times N)}$ is the called the output matrix and $v_k \in R^{(M \times 1)}$ is a sequence of *iid* $N(0, R_k)$ scalar random variables called the measurement noise. We assume that R_k (which is a non-negative symmetric matrix) is non-singular. Let $\mathcal{Y}_k := \{y_0, y_1, \ldots, y_k\}$ denote the measurements up to time k. The symbol $:=$ denotes "defined as".

2.1 The Filtering Problem

The filtering problem stated in the broadest terms is to determine information about the state x_k from measurements up until time k (or $k-1$). For stochastic processes, this information is represented by the conditional probability density functions (PDFs) $p(x_k|\mathcal{Y}_k)$ (or $p(x_k|\mathcal{Y}_{k-1})$). Usually we limit our interest to statistics of the state such as the mean and variance that can be calculated from these PDFs, but this is not always the case (see Section 6 later in this report). The filtering problem then becomes the problem of estimating these particular statistics.

For the special case of the linear Gauss-Markov system with Gaussian initial conditions it can be shown that the PDFs at each time instant are Gaussian and hence the PDFs can be completely specified by the mean and variance statistics (which is a finite number of quantities). For this special case, a finite-dimensional filter (the Kalman filter) can be implemented that optimally calculates these two statistics and hence completely specifies the PDFs at each time instant.

In general terms, the linear Gauss-Markov filtering problem is a special case and for most other filtering problems the relevant PDFs can not be represented by a finite number of statistics and finite-dimensional optimal filters are not possible.

Let us now concentrate on the filtering problem associated with the Kalman filter. Let us define the conditional mean estimates $\hat{x}_{k|k-1} := E[x_k|\mathcal{Y}_{k-1}]$ and $\hat{x}_{k|k} := E[x_k|\mathcal{Y}_k]$. And let us also define the error covariance matrices $P_{k|k-1} := E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})'|\mathcal{Y}_{k-1}]$ and $P_{k|k} := E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})'|\mathcal{Y}_k]$

The discrete-time Kalman filtering problem is stated as follows: *For the linear system (2.1)(2.2) defined for $k \geq 0$ for which the initial state x_0 is a Gaussian random variable with mean $\hat{x}_0$ and covariance P_0 which are independent of $\{w_k\}$ and $\{v_k\}$, determine the estimates $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$ and the associated error covariance matrices $P_{k|k-1}$ and $P_{k|k}$.*

2.2 The Kalman Filter

It is well known that the discrete-time Kalman filter is a finite-dimensional optimal solution to this problem [1, 2, 3]. Here finite-dimensional in the sense that it can be implemented exactly with a finite number of operations and a finite amount of memory and optimality is in a conditional mean sense. The optimality of the Kalman filter can be established in a variety of ways and we refer the reader to [1] for a good presentation. In this report we simply present the Kalman filter without proof.

The standard implementation of the Kalman filter is the following recursion.

$$
\begin{aligned}
\hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} \\
P_{k|k-1} &= A_{k-1}P_{k-1|k-1}A'_{k-1} + B_{k-1}Q_kB'_{k-1} \\
K_k &= P_{k|k-1}C'_k\left[C_kP_{k|k-1}C'_k + R_k\right]^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\left[y_k - C_k\hat{x}_{k|k-1}\right]
\end{aligned}
$$

$$P_{k|k} = P_{k|k-1} - K_k C_k P_{k|k-1} \tag{2.3}$$

where $\hat{x}_{0|0} = E[x_0] = \hat{x}_0$ and $P_{0|0} = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)'] = P_0$. The recursion (2.3) is one of the many algebraically equivalent recursions for the same estimates $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$.

It is common to interpret each iteration of the Kalman filter equations as two steps. The first two equations in (2.3) correspond to a time update process and the final three equations as a measurement update process. In the time update stage, the previous estimate is used to predict the state value (and the covariance) at the next time instant. In the measurement update stage, the prediction of the state (and the covariance) is corrected using the information in the new measurement.

Alternatively, the Kalman filtering problem can be interpreted as designing K_k such that the filter error is white (in particular uncorrelated), see Figure 1. To design K_k, knowledge of A_k, C_k, Q_k and R_k is required. If these quantities are not known with complete certainty then the conditional mean optimal filter can not be designed.

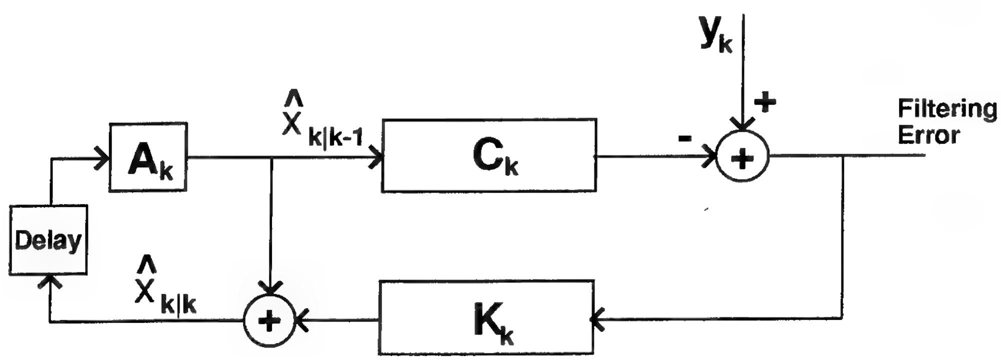

Figure 1 (U): Block diagram for state estimate update as a whitening filtering

Remarks

1. It can be shown that the Kalman filter: is the minimum variance filter [1, 8], is unbiased [1], and is a recursive Bayesian estimator optimal in a minimum mean square sense. In addition, if the Gaussianity assumptions on x_0, $\{w_k\}$ and $\{v_k\}$ no longer hold then the Kalman filter is the minimum variance filter among the class of linear filters [1, 8].

2. Under reasonable conditions (see [1, 8]) the Kalman filter is asymptotically stable in the sense that it will exponentially forget any initial condition errors. This is an important property because it means that errors introduced during filtering do not necessarily make the Kalman filter diverge.

3. The recursion (2.3) is the standard form of the Kalman Filter. Another common form of the Kalman filter is the innovations form [1]. There are many other forms of the filter that include slight modifications of the covariance and gain equations and in the next section we consider one form that has improved numeric stability.

3 U-D Factorisation Form of the Kalman Filter

3.1 Filter Stability

According to Maybeck [2]: "An algorithm can be said to be numerically stable if the computed result of the algorithm corresponds to an exactly computed solution to a problem that is only slightly perturbed from the original one." The standard formulation of the Kalman filter is not always numerically stable in this sense [4].

There can be numerical problems in implementing the above Kalman filter recursions on a finite word-length digital processor. For example, although it is theoretically impossible for the covariance matrices to have negative eigenvalues, such a situation can arise due to inaccuracies in numerical calculations. Such a condition can lead to instability or divergence of the Kalman filter.

To avoid the numerical problems of the standard formulation of the Kalman filter equations, many alternative recursions have been developed [2, 1]. These alternative recursions update the state and covariance matrices in a way that ensures that numerical problems are avoided. These formulations are algebraically equivalent to the standard Kalman filter equations but exhibit improved numerical precision and stability [2]. Because these alternative forms are algebraically equivalent, the design and tuning of the optimal filter can be done using the standard equations, ignoring errors caused by numerical instability, and the numerically stable forms only need be considered when the actual implementation is performed [2].

It should be noted that the sort of numerical instabilities inherent in the standard Kalman filter equations are issues that appear when some of the measurements are very accurate [1], when the process covariance has large dynamic range [7] or when the Kalman filter is applied to non-linear problems through a problem formulation such as the extended Kalman filter. In general terms, the numerical precision of modern computational technologies has reduced the need to use numerically stable forms of the Kalman filter.

There are two major advantages in using one of the numerical stable forms of the Kalman filter [1]:

1. The recursions can be formulated in a way that ensure that computational errors cannot lead to a matrix ($P_{k|k}$ or $P_{k|k-1}$) that fails to be nonnegative definite or symmetric.

2. The numerical conditioning of the recursions can be improved so that only half as much precision is required in the computations.

The approach presented here to enhance the numerical characteristics of the Kalman filter is known as the "U-D covariance factorisation" developed by Bierman and Thornton [4, 5, 6]. There are several other forms of the Kalman filter with increased numeric stability, such as the square root forms, see [2, 1] for details. We consider only the U-D factorisation in this report because it offers a reasonable compromise of numerical stability and increased computational effort [2].

3.2 The U-D Factorisation of the Kalman filter

The U-D factorisation involves expressing the error covariance matrices as

$$\begin{aligned} P_{k|k-1} &= U(k|k-1)D(k|k-1)U(k|k-1)' \quad \text{and} \\ P_{k|k} &= U(k|k)D(k|k)U(k|k)' \end{aligned} \tag{3.1}$$

where U are unitary upper triangular matrices and D are diagonal. The above U-D factorisation always exists for square symmetric, positive semi-definite matrices [2]; hence, the error covariance matrices $P_{k|k}$ and $P_{k|k-1}$ can always be factorised this way.

Although the U-D factorisation is not unique we can work with the following factorisation for initialisation purposes: For a P (either $P_{k|k-1}$ or $P_{k|k}$) ,

$$\begin{aligned} D_{jj} &= P_{jj} - \sum_{\ell=j+1}^{N} D_{\ell\ell} U_{j\ell}^2 \quad \text{and} \\ U_{ij} &= \begin{cases} 0 & i > j \\ 1 & i = j \\ (P_{ij} - \sum_{\ell=j+1}^{N} D_{\ell\ell} U_{i\ell} U_{j\ell})/D_{jj} & i = j-1, j-2, \ldots, 1 \end{cases} \end{aligned} \tag{3.2}$$

where U_{ij} is the ij th element of U etc. and

$$\begin{aligned} D_{NN} &= P_{NN} \quad \text{and} \\ U_{iN} &= \begin{cases} 1 & i = N \\ P_{iN}/D_{NN} & i = N-1, N-2, \ldots, 1. \end{cases} \end{aligned} \tag{3.3}$$

To develop the filter itself we consider the time update and measurement update separately in the following subsections.

3.3 Time Update

It can be shown that the time update equations for the U-D factorisation of the Kalman filter can be implemented as follows (taken from [2]). Using the following definitions:

$$\begin{aligned} Y_k &:= [A_{k-1} U^{(k-1|k-1)} \big| B_{k-1}], \\ \tilde{D}_{k+1} &:= \begin{bmatrix} D^{(k-1|k-1)} & \mathbf{0} \\ \mathbf{0} & Q_{k-1} \end{bmatrix}, \\ [a_1 | a_2 | \ldots | a_N] &:= Y_k' \quad \text{and} \\ U_{jj}^{(k|k-1)} &:= 1 \text{ for } j = 1, \ldots, N \end{aligned} \tag{3.4}$$

where $\mathbf{0}$ are zero matrices, $D^{(k-1|k-1)}$ and $U^{(k-1|k-1)}$ denote the U-D factorisation of $P_{k-1|k-1}$, Y_k is an $(N \times (N+S))$ matrix, $\tilde{D}_{k+1}$ is an $(N+S \times N+S)$ matrix and a_i are $2N$ vectors. Then the time update equations can be calculated recursively for $\ell = N, N-1, \ldots, 1$:

$$\begin{aligned} c_\ell &= \tilde{D}_k a_\ell \\ D_{\ell\ell}^{(k|k-1)} &= a_\ell' c_\ell \\ d_\ell &= c_\ell / D_{\ell\ell}^{(k|k-1)} \\ \left.\begin{aligned} U_{j\ell}^{(k|k-1)} &= a_j' d_\ell \\ a_j &= a_j - U_{j\ell}^{(k|k-1)} a_\ell \end{aligned}\right\} & \quad j = 1, 2, \ldots, \ell-1 \end{aligned} \tag{3.5}$$

where $D^{(k|k-1)}$ and $U^{(k|k-1)}$ denote the U-D factorisation of $P_{k|k-1}$. Here a_ℓ, c_ℓ and d_ℓ are temporary variables that are reused. The state estimate is updated as follows:

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1}. \tag{3.6}$$

3.4 Measurement Update

The scalar measurement update equations for the U-D factorisation form of the Kalman filter can calculated using the following definitions:

$$\begin{aligned}
f &:= U^{(k|k-1)\prime}C_k' \\
v_j &:= D_{jj}^{(k|k-1)}f_j \qquad \text{for } j = 1, 2, \ldots, N \\
a_0 &= R \qquad \text{and} \qquad (3.7) \\
U_{jj}^{(k|k)} &:= 1 \qquad \text{for } j = 1, 2, \ldots, N. \qquad (3.8)
\end{aligned}$$

Then the time update equations can be calculated recursively for $\ell = 1, \ldots, N$:

$$\begin{aligned}
a_\ell &= a_{\ell-1} + f_\ell v_\ell \\
D_{\ell\ell}(k|k) &= D_{\ell\ell}(k|k-1)a_{\ell-1}/a_\ell \\
b_\ell &= v_\ell \\
p_\ell &= -f_\ell/a_{\ell-1} \\
\left.\begin{aligned} U_{j\ell}(k|k) &= U_{j\ell}(k|k-1) + b_j p_\ell \\ b_j &= b_j + U_{j\ell}(k|k-1)v_\ell \end{aligned}\right\} & \quad j = 1, 2, \ldots, (\ell - 1) \qquad (3.9)
\end{aligned}$$

where a_k, b_k and p_k are temporary variables that are reused. Vector measurement updates can be handled as a number of scalar measurement updates.

Let b denote the vector formed from elements b_ℓ, then the state is updated as follows:

$$\begin{aligned}
K_k &= b/a_N \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}). \qquad (3.10)
\end{aligned}$$

3.5 Computational Effort

The improved numeric properties of the U-D factorisation form of the Kalman filter are achieved at the cost of increased computational effort. According to Maybeck [2] the

computational effort required to implement the U-D factorisation is slightly greater than the standard form. However, considering the requirement for numerical stability, the U-D form of the algorithm appears to offer a reasonable compromise between computational efficiency and numerical stability. See [2] for more details and a description of the actually computational effort required.

In most situations the standard Kalman filter can be made to work by increasing the precision of computations as required or using ad hoc modifications to reduce the dynamic range of variables [1]. However, in these situations, better performance can often be achieved by the U-D form of the Kalman filter using less precision in the computations. For these reasons, a number of practitioners have argued that numerically stable forms of the Kalman filter should always be used in preference to the standard form of the Kalman filter, rather than switching to a stable form when required [2].

Finally, it should be noted that in high measurement noise situations the standard Kalman filtering equations tend to be fairly stable, reducing the motivation for numerically stable forms. In fact, in some high measurement noise situations, the standard Kalman filter equations can be simplified though approximations that considerably reduce computational burden with little loss in optimality and without introducing stability problems [1].

4 Extended Kalman Filter

The discussion so far in this report has been directed towards optimal filtering of linear systems using the Kalman filter. However, in most realistic applications the underlying physical system is non-linear. In some situations, slightly non-linear systems can be approximated as linear systems and the Kalman filter provides a satisfactory filtering solution. In other situations, the system may have obvious non-linear characteristics that can not be ignored.

Filtering for non-linear systems is a difficult problem for which few satisfactory solutions can be found (we consider several algorithms in this report that are somewhat satisfactory in some situations). The sub-optimal approach considered in this section is an extension of the Kalman filter known as the extended Kalman filter.

The extended Kalman filter is posed by linearisation of a non-linear system. Consider the following non-linear system defined for $k \in Z^+$:

$$\begin{aligned} x_{k+1} &= a_k(x_k) + b_k(x_k)w_k \\ y_k &= c_k(x_k) + v_k \end{aligned} \tag{4.1}$$

where $a_k(.)$, $b_k(.)$ and $c_k(.)$ are non-linear functions of the state and w_k,v_k are defined as before. Let us define the following quantities:

$$A_k = \left.\frac{\partial a_k(x)}{\partial x}\right|_{x=\bar{x}_{k|k-1}}, \quad B_k = \left.\frac{\partial b_k(x)}{\partial x}\right|_{x=\bar{x}_{k|k-1}} \quad \text{and} \quad C_k = \left.\frac{\partial c_k(x)}{\partial x}\right|_{x=\bar{x}_{k|k-1}}. \tag{4.2}$$

Let us also introduce matrices Q_k^* and R_k^* which are related to the covariance matrices for noises w_k and v_k. However, as will be shown later in Section 4.2, the matrices Q_k^* and R_k^* need not equal Q_k and R_k and other positive definite matrices are often better choices.

The extended Kalman filter is implemented using the following equations:

$$\begin{aligned} \bar{x}_{k|k-1} &= a_{k-1}(\bar{x}_{k-1|k-1}) \\ \bar{P}_{k|k-1} &= A_{k-1}\bar{P}_{k-1|k-1}A'_{k-1} + B_{k-1}Q_k^*B'_{k-1} \\ K_k &= \bar{P}_{k|k-1}C'_k\left[C_k\bar{P}_{k|k-1}C'_k + R_k^*\right]^{-1} \\ \bar{x}_{k|k} &= \bar{x}_{k|k-1} + K_k\left[y_k - c_k(\bar{x}_{k|k-1})\right] \\ \bar{P}_{k|k} &= \bar{P}_{k|k-1} - K_kC_k\bar{P}_{k|k-1}. \end{aligned} \tag{4.3}$$

The equations are no longer optimal or linear because A_k etc. depend on $\bar{x}_{k|k-1}$ etc. The symbols $\bar{x}_{k|k-1}$, $\bar{x}_{k-1|k-1}$, $\bar{P}_{k|k-1}$ and $\bar{P}_{k-1|k-1}$ now loosely denote approximate conditional means and covariances respectively.

The extended Kalman filter presented above is based on first order linearisation of a non-linear system, but there are many variations on the extended Kalman filter based on second order linearisation or iterative techniques. Although the extended Kalman filter or other linearisation techniques are no longer optimal, these filters can provide reasonable filtering performance in some situations.

4.1 U-D Covariance Form of the Extended Kalman Filter

A U-D factorisation form of the extended Kalman filter can be posed by appropriate modification of U-D factorisation of the Kalman filter as follows.

Equations (3.10) and (3.6) are replaced by the following equations:

$$\begin{aligned}\bar{x}_{k|k-1} &= a_{k-1}(\bar{x}_{k-1|k-1}) \quad \text{and} \\ \bar{x}_{k|k} &= \bar{x}_{k|k-1} + K_k(y_k - c_k(\bar{x}_{k|k-1})).\end{aligned} \tag{4.4}$$

4.2 Stochastic Stability of the Extended Kalman Filter

A key question when applying an extended Kalman filter to a particular non-linear problem is when will the extended Kalman filter be stable and when will it diverge? Heuristic arguments have been used to suggest that if the non-linearities are linear enough and the filter is initialised well enough then the filter should be stable. This heuristic argument has encouraged the use of the extended Kalman filter in a wide variety of signal processing, control and filtering problems. However, without any solid stability results, the error behaviour of the extended Kalman filter needs to be examined through testing whenever applied [8, 1].

Recently, solid stability results have established conditions on the non-linearities and initial conditions which ensure that the extended Kalman filter will produce estimates with bounded error [9, 10]. These results answer some of the stability questions surrounding the extended Kalman filter [9, 10]. This section repeats the stability results of [10].

Consider again the non-linear system (4.1) defined in Section 4:

$$\begin{aligned}x_{k+1} &= a_k(x_k) + b_k(x_k)w_k \\ y_k &= c_k(x_k) + v_k.\end{aligned} \tag{4.5}$$

Let us define the following quantities

$$\begin{aligned}\varphi(x_k, \bar{x}_k) &:= a_k(x_k) - a_k(\bar{x}_k) - A_k(x_k - \bar{x}_k) \\ \chi(x_k, \bar{x}_k) &:= c_k(x_k) - c_k(\bar{x}_k) - C_k(x_k - \bar{x}_k)\end{aligned}$$

where $\bar{x}_k$ is some estimate of the state (see Figure 2 for a graphic interpretation of $\varphi(x_k, \bar{x}_k)$). Also, we define the estimation error as $\tilde{x}_{k|k} := x_k - \bar{x}_{k|k}$. Then the following theorem is presented in [10].

Theorem 1 *(Theorem 3.1) Consider the nonlinear system (4.1) and the extended Kalman filter presented in Section 4. Let the following hold:*

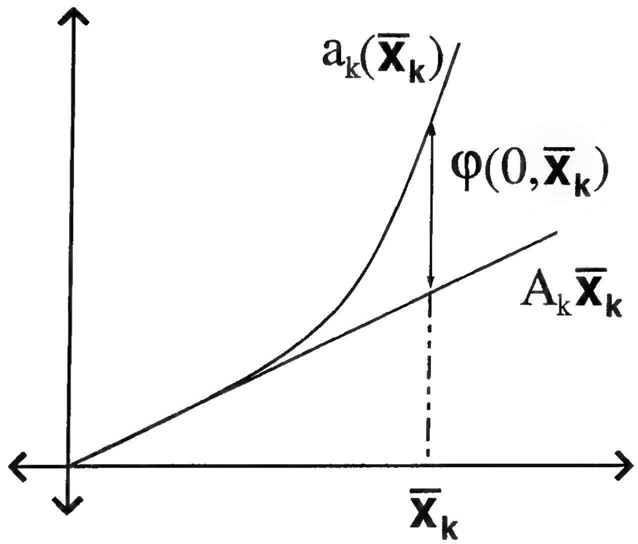

Figure 2 (U): Graphical interpretation of $\varphi(x_k, \bar{x}_k)$.

1. *There are positive real numbers* $\bar{a}, \bar{c}, \underline{p}, \bar{p}, \underline{q}, \underline{r} > 0$ *such that the following bounds hold for all* $k \geq 0$*:*

$$\begin{aligned} \|A_k\| &\leq \bar{a} \\ \|C_k\| &\leq \bar{c} \\ \underline{p}I &\leq P_{k|k-1} \leq \bar{p}I \\ \underline{q} &\leq Q_k^* \\ \underline{r} &\leq R_k^*. \end{aligned} \tag{4.6}$$

2. A_k *is nonsingular for all* $k \geq 0$.

3. *There are positive numbers* $\epsilon_\varphi, \epsilon_\chi, \kappa_\varphi, \kappa_\chi > 0$ *such that*

$$\begin{aligned} \|\varphi(x_k, \bar{x}_k)\| &\leq \kappa_\varphi \|x_k - \bar{x}_k\|^2 \\ \|\chi(x_k, \bar{x}_k)\| &\leq \kappa_\chi \|x_k - \bar{x}_k\|^2 \end{aligned} \tag{4.7}$$

for x_k, $\bar{x}_k$ *with* $\|x_k - \bar{x}_k\| \leq \epsilon_\varphi$ *and* $\|x_k - \bar{x}_k\| \leq \epsilon_\chi$ *respectively for all* k *where* $\bar{x}_k$ *is any estimate of* x_k *at time* k.

Then the estimation error is bounded with probability one, provided the initial estimation error satisfies

$$||\tilde{x}_o|| \leq \epsilon \tag{4.8}$$

and the covariances of the noise terms are bounded via $Q_k \leq \delta I$ and $R_k \leq \delta I$ for some δ, ϵ.

Remark

4. The proof of Theorem 1 is given in [10].

5. This result states that if the non-linearity is small then the EKF is stable if initialised close enough to the true initial value. The greater the deviation from linearity the better the initialisation needs to be.

The proof presented in [10] provides a technique for calculating conservative bounds for ϵ and δ. Simulation studies suggest that ϵ and δ can be significantly larger than these bounds in some situations.

We define the following to repeat the bounds presented in [10]

$$\bar{\epsilon} := \min(\epsilon_\varphi, \epsilon_\chi) \tag{4.9}$$

$$\bar{\kappa} := \kappa_\varphi + \bar{a}\bar{p}\frac{\bar{c}}{\underline{r}}\kappa_\chi. \tag{4.10}$$

We also define the following

$$\kappa_{nonl} := \frac{\bar{\kappa}}{\underline{p}}\left(2\left(\bar{a} + \bar{a}\bar{p}\frac{\bar{c}^2}{\underline{r}}\right) + \bar{\kappa}\bar{\epsilon}\right) \tag{4.11}$$

$$\kappa_{noise} := \frac{S}{\underline{p}} + \frac{\bar{a}^2\bar{c}^2\bar{p}^2}{\underline{p}\underline{r}^2} \tag{4.12}$$

$$\alpha := 1 - 1/\left(1 + \frac{\underline{q}}{\bar{p}\bar{a}^2(1+\bar{p}\bar{c}^2/\underline{r})^2}\right). \tag{4.13}$$

Then

$$\epsilon = \min\left(\bar{\epsilon}, \frac{\alpha}{2\bar{p}\kappa_{nonl}}\right) \tag{4.14}$$

and

$$\delta = \frac{\alpha\epsilon^2}{2\bar{p}\kappa_{noise}}. \tag{4.15}$$

To gain an understanding of the stability of the extended Kalman filter consider the following special case.

Example: Quadratic Non-linearities

Consider the situation where system functions, $a_k(.)$ and $c_k(.)$ are quadratic in the state. Then φ and χ are bounded in (4.7) for all ϵ_φ and ϵ_χ. Hence κ_{nonl} is unbounded and hence ϵ is unbounded. The conclusion is that if the non-linear elements in the model are quadratic then the extended Kalman filter is stable for any initial estimate.

4.3 Higher-Order Approximation Methods and Other Issues

The extended Kalman filter presented above is based on a first-order Taylor series approximation of the system non-linearity. There are a number of variations and generalisation of the ideas used in the extended Kalman filter [1, p.196]. These ideas include second (and higher order) Taylor series approximation of the non-linear system, banks of extended Kalman filters (or the Gaussian sum approach), Monte-Carlo simulation techniques [27] (related to the particle filter ideas presented in Section 6), higher-order moment filters, and others.

In general terms, any one of these filtering approaches may be superior to the extended Kalman filter in a particular situation, but none of these approaches is superior in all applications. The particular type of non-linearities present in a system may make particular approaches more fruitful than others, but there are no real guidelines. Some approaches, such as Monte-Carlo techniques, may appear to be suitable to a large class of problems; however, their general applicability is at the cost of elegance and computational load.

A reasonable approach (based on the previous experiences of the author) maybe to use the simplest feasible filtering approach and then re-evaluate the choice of filter if the desired performance is not achieved. It is likely that the computational requirements of one of the more complicated approximate filtering techniques is sufficient motivation to first investigate a simpler approach.

One possible trap to avoid is the posing of unrealistic filtering problems. It is easy to pose filtering problems that are not solvable due to observability problems. In other situations, reasonable filtering performance can be difficult to obtain because the observed system is close to unobservable. Observability is a control system's concept that is defined as the

ability to determine the state from observations. For linear systems there are easy tests for observability (see [1, 29] or an undergraduate control systems text) but questions of observability are more complicated in non-linear systems [28].

The observability condition is only one part of the question, because a system is either observable or not. A practitioner should also be interested in the relative difficulty of a filtering problem. One way of quantifying the difficulty of achieving good state estimates in a particular filtering problem is via the Fisher information matrix and the Cramer-Rao lower bound [31, 30].

The Fisher information matrix describes how much information about the state variable is available in the measurements. The Cramer-Rao bound is a lower bound on the amount of data required in a particular problem to achieve a particular amount of certainty in estimation. A Fisher information matrix with poor characteristics highlights that a particular problem is difficult in a fundamental way (and changing the filtering algorithm will probably not significantly improve filtering performance). When a system is not observable or the Fisher information matrix for the problem has poor characteristics, the practitioner will probably need to consider redesigning the system.

The task of choosing an appropriate model and filtering algorithm for a non-linear system is non-trivial. It is implicit in the preceding discussion that the system model must reflect the nature of the true system. This report does not address any of the approaches (neither data based nor first principles based) for designing models of systems, see [30] for an introduction. Once a system model has been obtained, the observability of this system should be tested and the Fisher information matrix examined to determine the difficulty of the filtering problem.

Once comfortable that the measurements have enough information, a filtering algorithm can be chosen. It is difficult to give general guidelines for algorithm choice but experience with similar problems can provide useful insights. A practitioner will probably have a bias towards their favorite algorithm, or an algorithm that worked on a similar problem, and may try this algorithm first.

Quantifying the achieved performance of an algorithm can also be difficult. Comparison with similar problems can both be helpful and misleading. A reasonable approach would be to compare the performance of the algorithm with the performance of an higher-order

algorithm. This comparison may give a feel for the trade-off between complexity and filtering performance for this problem and what sort of algorithm will give the desired filtering performance.

In the next section we consider the filtering when the system model is not known with complete certainty.

5 Robust Filtering

The Kalman filter has been applied to a large class of problems. However, in many situations the system model is non-linear or uncertain in the sense that the model is not completely known. The previous section dealt with use of the extended Kalman filter on a linearised version of a non-linear system. Even in situations where the non-linear system is completely known the extended Kalman filter may give poor performance. In a situation where there is model uncertainty, it can be shown that the Kalman filter and the extended Kalman filter are often very poor filtering choices [24].

In recent years, a number of new approaches have been proposed for the filtering problem that are related to robust control techniques. This includes approaches such as H^∞ filtering [25], risk-sensitive filtering [11, 12] and robust filtering [24].

5.1 The Robust Kalman Filter

In this section we describe the development of the robust Kalman filter for uncertain discrete-time systems (this development is given in [24]). The description given here is in terms of a uncertain systems where the uncertainty has a particular form. However, the results can be used for fairly general systems by transforming the uncertain system into the following form.

Consider the following time-varying uncertain discrete-time system defined for $k \in Z^+$:

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k^1 w_k, \\ z_k &= K_k x_k, \\ y_k &= C_k x_k + v_k \end{aligned} \tag{5.1}$$

where again x_k is the state, $w_k \in R^p$ and $v_k \in R^\ell$ are uncertainty inputs, $z_k \in R^q$ is the uncertainty output, $y_k \in R^\ell$ is the measured output, and A_k, B_k^1, K_k and C_k are given matrices such that A_k is non-singular.

The uncertainty in this system is constrained to satisfy the following constraint for all k.

Let $N = N' > 0$ be a given matrix, $\bar{x}_0 \in R^n$ be a given vector, $d < 0$ be a given constant, and Q_k and R_k be given positive-definite symmetric matrices. The following constraint is termed the sum quadratic constraint:

$$(x_0 - \bar{x}_0)'N(x_0 - \bar{x}_0) + \sum_{\ell=0}^{k-1}(w_\ell' Q_\ell w_\ell + v_{\ell+1}' R_{\ell+1} v_{\ell+1}) \leq d + \sum_{\ell=0}^{k-1} ||z_{\ell+1}||^2 \quad \text{for all } k. \quad (5.2)$$

A system (5.1) satisfying the sum quadratic constraint can be thought of as a linear system driven by a process noise and observed in a measurement noise which can be anywhere in a constrained set. This description of an uncertain system is fairly general because uncertainties in system matrices such as A can be thought of as noise terms.

Before proceeding to present filtering results, we give an example of a structured uncertainty system that satisfies (5.1) and (5.2). Consider a system where there is some uncertainty in the state transition matrix. The uncertainty could result because the dynamics of the physical system can not be determined but these dynamics need to be modelled mathematically in some way. This uncertainty may be represented by the following equations (this example is given in [24]):

$$\begin{aligned} x_{k+1} &= [A_k + B_k^{11}\Delta_k K_k]x_k + B_k^{21} n_k, \\ y_k &= C_k x_k + \bar{n}_k, \quad \text{with} \quad ||\Delta_k' Q_k^{\frac{1}{2}}|| \leq 1 \end{aligned} \quad (5.3)$$

where Δ_k is the uncertainty matrix, n_k and $\bar{n}_k$ are noise sequences, $B_k^1 = [B_k^{11}, B_k^{12}]$, and $||.||$ denotes the standard induced matrix norm. The system uncertainty is in the state transition matrix because the value of Δ_k is not known. Also, let this system satisfy the condition

$$(x_0 - \bar{x}_0)'N(x_0 - \bar{x}_0) + x_0' K_0' Q_0 K_0 x_0 + \sum_{\ell=0}^{k-1}(n_\ell)' Q_\ell n_k + \sum_{\ell=0}^{k}(\bar{n}_k)' R_k \bar{n}_k \leq d.$$

To establish that (5.3) is admissible for the uncertain system described by (5.1),(5.2), let

$$w_k = \begin{bmatrix} \Delta_k K_k x_k \\ n_k \end{bmatrix}$$

and $v_k = \bar{n}_k$ for all k where

$$||\Delta_k' Q_k^{\frac{1}{2}}|| \leq 1$$

for all k. Then condition (5.2) is satisfied. Hence system (5.3) satisfies the requirements of the theory and the results developed in the following section can be applied to this system.

5.2 The Set-value State Estimation Problem

The filtering problem examined in this section can be stated as follows: given an output sequence $\{y_0, y_1, \ldots, y_k\}$ then find the corresponding set of all possible states x_k at time k with uncertainty inputs and initial conditions satisfying the constraint (5.2).

Definition 1 *The system (5.1),(5.2) is said to be* **strictly verifiable** *if the set of possible states x_k at time k is bounded for any x_0, $\{y_k\}$ and d.*

Definition 2 *The output sequence $\{y_0, \ldots, y_k\}$ is* **realizable** *if there exist sequences $\{x_k\}$, $\{w_k\}$ and $\{v_k\}$ satisfying (5.1) and (5.2).*

In addition to solving the state estimation problem, the results that follow also solve the following problem: given an output sequence, determine if this output is realizable for the uncertainty system [24]. Thus, the following results are useful in answering questions of model validation [24].

The solution to the filtering problem involves the following Riccati difference equation:

$$\begin{aligned} F_{k+1} &= [\bar{B}_k' S_k \bar{B}_k + Q_k]^{\#} \bar{B}_k' S_k \bar{A}_k, \\ S_{k+1} &= \bar{A}_k' S_k [\bar{A}_k - \bar{B}_k F_{k+1}] + C_{k+1}' R_{k+1} C_{k+1} - K_{k+1}' K_{k+1} \\ S_0 &= N \end{aligned} \tag{5.4}$$

where

$$\bar{A}_k := A_k^{-1}, \quad \bar{B}_k := \bar{A}_k B_k^1$$

and $[.]^{\#}$ denotes the Moore-Penrose pseudo-inverse (see [1]) if an inverse does not exist.

Solutions to the Riccati equations are required to satisfy the following conditions:

$$\begin{aligned} \bar{B}_k' S_k \bar{B}_k + Q_k &\geq 0 \quad \text{and} \\ \mathcal{N}(\bar{B}_k' S_k \bar{B}_k + Q_k) &\subset \mathcal{N}(\bar{A}_k' S_k \bar{B}_k) \end{aligned} \tag{5.5}$$

for all k. Here $\mathcal{N}$ denotes the operation of taking the null space of a matrix.

The state estimate solution requires the following additional equations [24].

$$\begin{aligned}
\eta_{k+1} &= [\bar{A}_k - \bar{B}_k F_{k+1}]'\eta_k + C'_{k+1} R_{k+1} y_{k+1}, \\
\eta_0 &= N x_0, \\
g_{k+1} &= g_k + y'_{k+1} R_{k+1} y_{k+1} - \eta'_k \bar{B}_k [\bar{B}'_k S_k \bar{B}_k + Q_k]^{\#} \bar{B}'_k \eta_k, \\
g_0 &= x'_0 N x_0.
\end{aligned} \tag{5.6}$$

Theorem 2 *[24] Consider the uncertainty system (5.1),(5.2). Then the following statements hold:*

(i) The uncertainty system (5.1),(5.2) is strictly verifiable if and only if there exists a solution to the Riccati equation (5.4) satisfying condition (5.5).

(ii) Suppose the uncertainty system (5.1),(5.2) is strictly verifiable. Then the output sequence $\{y_k\}$ is realizable if and only if $\rho_k(\{y_k\}) \geq -d$ where

$$\rho_k(\{y_k\}) := \eta'_k S_k^{-1} \eta_k - g_k.$$

(iii) Suppose the uncertainty system (5.1),(5.2) is strictly verifiable, then the set of possible state values at time k is

$$\left\{ x_k \in R^n : \left\| S_k^{\frac{1}{2}} x_k - S_k^{-\frac{1}{2}} \eta_k \right\|^2 \leq p_k(\{y_k\}) + d \right\}. \tag{5.7}$$

Proof: See [24, pp. 75-77].

The center of the solution set (5.7) can be used as a state estimate and hence it follows that the state estimate at time k is $\hat{x}_k = S_k^{-1} \eta_k$ (see [24]).

To demonstrate the performance of the filter, consider the following example.

5.3 An Example Robust Kalman Filtering Problem

This example was given in [24] to illustrate the robust filter. Consider the following structured uncertainty system, ie. the form of (5.3).

$$\begin{aligned}
\begin{bmatrix} x_{k+1}(1) \\ x_{k+1}(2) \\ x_{k+1}(3) \end{bmatrix} &= \begin{bmatrix} 1.98 + 0.0127\Delta_k & -1 & 0 \\ 1 & 0 & 0 \\ 0.4 & 0 & 0.2 \end{bmatrix} \begin{bmatrix} x_k(1) \\ x_k(2) \\ x_k(3) \end{bmatrix} + \begin{bmatrix} 0.707 \\ 0 \\ 0 \end{bmatrix} n_k, \\
y_k &= x_k(3) + \bar{n}_k
\end{aligned} \tag{5.8}$$

where $x_0 = [0\ 0\ 0]'$,

$$(10 + 0.0127^2)||x_0||^2 + \sum_{\ell=0}^{k-1}(n_\ell)^2 + \sum_{\ell=1}^{k}(\bar{n}_\ell)^2 \leq 1$$

and the uncertainty parameter Δ_k satisfies

$$||\Delta_k|| \leq 1$$

for all k. Note that even if $\Delta_k = 0$ this filtering problem does not reduce to the linear-Gauss filtering problem solved by the Kalman filter because the noise terms satisfy the above sum quadratic constraint rather than having Gaussian density functions.

In this problem, the state equation is known with some, but not complete, certainty. The variable Δ_k parameterises the possible values of the state transition matrix. We know only that this parameter is constrained to a particular range.

To apply the results of Theorem 2 to this filtering problem, we consider the system in the uncertainty form (5.1). In this case, the matrices A, N, B^1, x_0, K, C, Q, and R are given by

$$\begin{aligned}
&A = \begin{bmatrix} 1.98 & -1 & 0 \\ 1 & 0 & 0 \\ 0.4 & 0 & 0.2 \end{bmatrix}, N = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}, B^1 = \begin{bmatrix} 0.707 & 0.707 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, x_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
&C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad K = \begin{bmatrix} 0.0127 & 0 & 0 \end{bmatrix}, \quad Q = I \text{ and } R = 1.
\end{aligned} \tag{5.9}$$

The constant d is given by $d = 1$.

To illustrate the performance of the robust filter, the system was subjected to the following noise sequences:

$$n_k = \begin{cases} 0.5 & \text{for } k = 0, \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad \bar{n}_k = \frac{1}{105}\sin(k/10).$$

It is straightforward to verify that the uncertainty input sequences

$$w_k = [\ x_k' K' \Delta' \quad (n_k)'\ ]' \quad \text{and} \quad v_k = \tilde{n}_k$$

satisfy the sum quadratic constraint (5.2).

Figure 3 shows estimates of the first component of the state from the robust Kalman filter applied to measurements from this uncertainty system. Lower and upper bounds for the possible values of the state have also been plotted. The bounds on the state estimate were obtained by numerically finding the largest value of the first component of the state in the solution set (5.7).

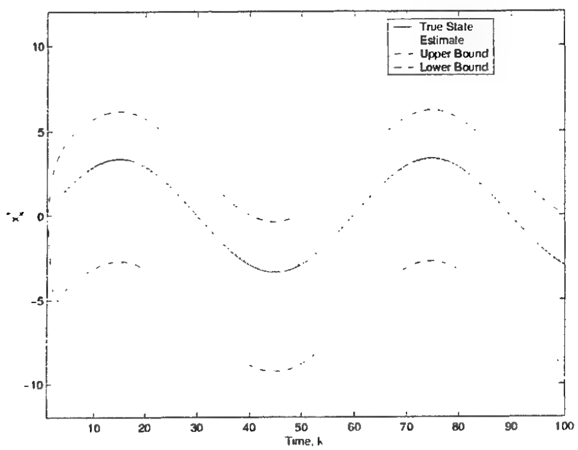

Figure 3 (U): Results of Robust Kalman Filter.

It is interesting to compare the performance of the robust Kalman filter and the standard Kalman filter, based on the nominal system (ie. with $\Delta_k = 0$), for this uncertainty system (see Figure 4). The performance of the standard Kalman filter is very similar to the robust Kalman filter and this suggests that there is no need for robust filtering with this system. Increasing the energy of the noise sequences does not result in a situation where the robust filter is significantly better than the standard Kalman filter.

6 Particle Filtering

An alternative approach to the filtering problem for general non-linear systems is the particle filter. The previous approaches in this report could be called parametric (or model based) techniques. Unlike the previous approaches, the particle filter attempts

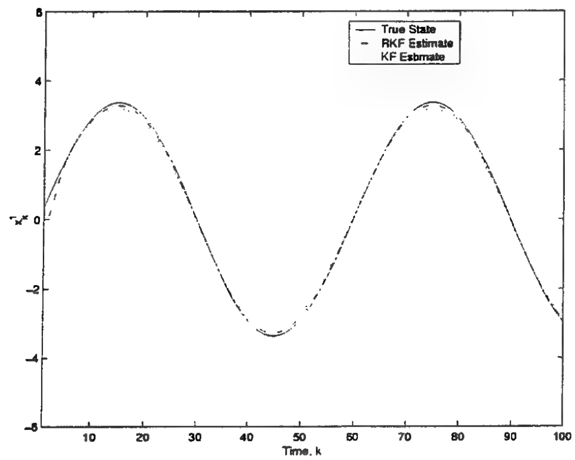

Figure 4 (U): Comparison of Robust Kalman Filter and the Kalman Filter.

to estimate the whole posterior PDF rather than particular statistics of the posterior PDF. The particle filter is a non-parametric technique because it does not attempt to parameterise the PDF.

The term particle filter includes the condensation algorithm, Bayesian bootstrap [16] or sampling importance resampling filter [15]. All these approaches represent the posterior probability density of the system state by a system of particles which evolve according to the non-linear system.

In general, a very large number of particles may be required to adequately represent the evolution of the system. However, it can be shown that any non-linear system can be approximated by a particle filter if the number of particles is large enough.

6.1 Non-linear System Model

We introduce a slightly more general non-linear model. Again x_k and y_k are the state process and observation process respectively. Then we define the system for $k \in Z^+$ as follows:

$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \quad \text{and} \\ y_k &= h_k(x_k, v_k) \end{aligned} \tag{6.1}$$

where w_k and v_k are noise terms with known distributions (not necessarily Gaussian). Given the initial PDF of the system, i.e. $p(x_0|y_{-1}) = p(x_0)$, these density functions can

in general be calculated for $k \geq 0$ as follows

$$
\begin{aligned}
p(x_k|\mathcal{Y}_{k-1}) &= \int p(x_k|x_{k-1})p(x_{k-1}|\mathcal{Y}_{k-1}) \quad \text{and} & (6.2)\\
p(x_k|\mathcal{Y}_k) &= \frac{p(y_k|x_k)p(x_k|\mathcal{Y}_{k-1})}{p(y_k|\mathcal{Y}_{k-1})} & (6.3)
\end{aligned}
$$

where

$$
p(y_k|\mathcal{Y}_{k-1}) = \int p(y_k|x_k)p(x_k|\mathcal{Y}_{k-1}). \tag{6.4}
$$

In the special case of linear dynamics and $p(x_0)$ having a Gaussian distribution then evaluation of the integrals can be simplified and this leads to the Kalman Filter. In general these integrals can not be analytically evaluated and the PDF must be approximated numerically. The particle filter technique is one way of developing a numeric approximation to the PDF.

6.2 Importance Sampling

The particle filter approach provides a Monte Carlo approximation to the PDF that consists of a set of random nodes in the state space s_k^i for $i = 1, \ldots, N^S$ (termed the support) and a set of associated weights, w_k^i for $i = 1, \ldots, N^S$, which sum to 1 (see Figure 5).

The objective of choosing the weights and support is to provide an approximation for the PDF such that

$$
\sum_{i=1}^{N^S} g(s_k^i)w_k^i \approx \int g(x_k)p(x_k)dx_k \tag{6.5}
$$

for typical functions g of the state space. This approximation is in the sense that the left-hand side equals the right-hand side as $N^S \to \infty$.

In importance sampling the support, s_k^i, is obtained by sampling values independently from a probability measure $\bar{p}(x_k)$ (termed the importance PDF) and attaching weights

$$
w_k^i = \frac{p(s_k^i)/\bar{p}(s_k^i)}{\sum_{j=1}^{N^S} p(s_k^j)/\bar{p}(s_k^j)}. \tag{6.6}
$$

Standard Importance Sampling Algorithm

The following algorithm is repeated from [15]. This algorithm uses the prior importance function as mentioned in [16]. This algorithm assumes that $p(x_{k+1}|x_k)$, $p(y_k|x_k)$ and $p(x_0)$ are known.

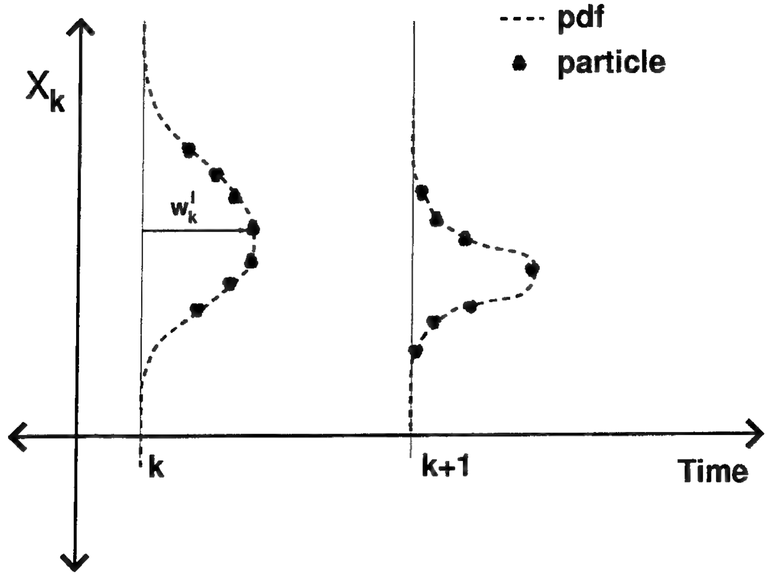

Figure 5 (U): Graphical interpretation of a particle filter

1. (Initialisation) Sample from $p(x_0)$ to obtain initial support s_1^i. Assign weights $w_1^i = 1/N^S$.

2. (Preliminaries on step k) Generate an approximation of any required statistics from the support and weight (s_{k-1}^i, w_{k-1}^i) approximation for $p(x_{k-1}|\mathcal{Y}_{k-1})$ using (6.5).

3. (Update using prior importance function) Now generate the next set of support points from the model dynamics, that is

$$s_k^i = f_{k-1}(s_{k-1}^i, w_{k-1}^i) \tag{6.7}$$

where w_{k-1}^i is sampled from the noise distribution and weights

$$w_k^i = \frac{p(y_k|s_k^i)}{\sum_{j=1}^{N^S} p(y_k|s_k^j)}. \tag{6.8}$$

4. Set k=k+1. Return to step 2

At any time, k expectations of $g(x_k)$ on $p(x_k|y_0, \ldots, y_k)$ can be estimated as follows

$$\int g(x_k)p(x_k|\mathcal{Y}_k)dx_k \approx \sum_{i=1}^{N^S} w_k^i g(s_k^i). \tag{6.9}$$

Hence, the mean and variance of x_k can be estimated as

$$\begin{aligned}\bar{x}_{k|k} &\approx \sum_{i=1}^{N^S} w_k^i (s_k^i) \\ \bar{P}_{k|k} &\approx \sum_{i=1}^{N^S} w_k^i (s_k^i - \bar{x}_{k|k})(s_k^i - \bar{x}_{k|k})'. \end{aligned} \tag{6.10}$$

A major limitation of the standard importance sampling is the degeneracy of the algorithm [16]. As k increases the support s_k^i tends to be concentrated on a same section of the state space (this is equivalent to having a smaller N^S). This effect is shown in Figure 5 where from time k to time $k+1$ the support has been concentrated towards the centre. It has been shown that the best approximation occurs when the variance of the weights is as small as possible [16]. The performance of the importance sampling algorithm can be improved by introducing the following resampling step

3.5 (Resample) The probability density function represented by (s_k^i, w_k^i) is resampled to an equally weighted support set $(s_k^i, (N^S)^{-1})$.

Remarks

6. The above algorithm uses the prior importance function which is very sensitive to outliers and in many situations an alternative importance function should be considered, see [16].

7. Further extensions including stratified sampling [15], the optimal importance function [16], linearised importance functions [16] and branching particle systems [17] may offer significant improvement over the importance sampling algorithm presented above.

8. Convergence proofs as $N^S \to \infty$ are given in [16] and [17].

9. Particle filter methods may require very large numbers of particles to ensure a reasonable approximation is obtained, see [17] for a low dimension example.

10. Particle filters can often be implemented in a parallel manner.

11. Particle filters can be sensitive to outliers [18] and the particle filter requires knowledge of the initial PDF much like the extended Kalman filter [15, 16].

6.3 Hidden Markov Model Approximation of the PDF

Another related method for approximating the evolution of the probability density function is to approximate the PDF as a grid of points which evolve by the state equations [20], see Figure 6. That is, the PDF is approximated by a probability mass function where the probability of being in particular regions of the state space are represented by a single number. Further, all points in the region are approximated by a single point in the region. Under this representation, the probability of being in a region of the state space is known, but it is not possible to determine the probability of being at a particular point in the region.

Under a few assumptions, the grid of points (or discretisation of the state space) can be considered a hidden Markov model to which the corresponding filtering theory can be applied, see [21] for filtering theory for HMMs. Although similar to the particle filter approach this method differs in that the grid of points represents regions of the state space (rather than fixed points) and the grid is fixed over time.

To develop a HMM approximation of the non-linear state equations we introduce the following notation. Let X_k denote the state of a discrete Markov state (or Markov chain) at time k. Let $\{R_1, \ldots, R_N\}$ denote N regions of the state space represented by the N states of the discrete Markov state. That is, if $x_k \in R_i$ then $X_k = i$. The evolution of a discrete Markov state is described by a state transition matrix A_k defined as follows:

$$A_k^{ij} := P(X_{k+1} = i | X_k = j)$$

where A_k^{ij} is the ijth element of $A_k \in R^{N \times N}$.

To approximate the evolution of x_k by a discrete state we approximate A_k as follows

$$A_k^{ij} = \frac{\int_{R_i} \int_{R_j} p(x_k | x_{k-1}) dx_{k-1} dx_k}{\int \int_{R_j} p(x_k | x_{k-1}) dx_{k-1} dx_k}.$$

It is assumed that $p(x_k | x_{k-1})$ is known and hence A_k defined this way can always be calculated. Once an initial value X_0 is given, the discrete Markov state approximation of the continuous state is completely specified.

Now an approximation of the non-linear observation process is required. The observation process $h_k(.,.)$ can be approximated be defining a matrix C_k as follows:

$$C_k^i = h_k(x^i, 0)$$

where C_k^i is the ith element of C_k and x^i is a representative value of x_k in the region R_i (for example the middle of the region). Then the observation process can be approximated as

$$y_k \approx C_k^{X_k} + v_k.$$

Standard HMM filtering solutions [21] on the HMM approximation can be applied if v_k is approximated as a Gaussian random variable. Let $\hat{X}_k^i$ denote $P(X_k = i | \{y_0, \ldots, y_k\})$. Then the HMM filter is

$$\hat{X}_k^i = \frac{1}{N_k} P(y_k | X_k = i) \sum_{j=1}^{N} A^{ij} \hat{X}_{k-1}^j \tag{6.11}$$

where

$$N_k = \sum_{i=1}^{N} P(y_k | X_k = i) \sum_{j=1}^{N} A^{ij} \hat{X}_{k-1}^j. \tag{6.12}$$

This filter can be used to estimate the probability of the state being in each region (or the PDF of the discretised system). Estimates of the state value or other statistics of the state can be formed from $\hat{X}_k$. For any general statistic, $E[g(x_k)|\{y_0, \ldots, y_k\}]$, the following estimate can be used:

$$E[g(x_k)|\{y_0, \ldots, y_k\}] \approx \sum_{i=1}^{N} g(x^i) \hat{X}_k^i. \tag{6.13}$$

In particular, the conditional mean estimate of the state is

$$\hat{x}_k \approx \sum_{i=1}^{N} x^i \hat{X}_k^i. \tag{6.14}$$

The HMM model that is developed is only an approximation of the true system but as the number of grid points increase the quality of the approximation should improve.

The advantage of applying HMM filters to the discretised state space system is that HMM filters are optimal in a conditional mean sense (on the discretised system) for any non-linear state equations [21]. If the discretised system represents the true system well, then the HMM filter should give good results. An additional advantage of the HMM filtering approach is that it does not suffer degeneracy in the same way as the particle filter. The grid is fixed in state space and hence the approximation to the PDF does not adapt as the nature of the PDF changes and hence degeneracy does not occur. Unfortunately, the fixed nature of the grid limits the accuracy of the approximation and the particle filter approach may perform better if the degeneracy can be controlled.

The major disadvantage of the HMM approach is that the state space can only be discretised to a fixed grid of points when the variation in state variable can be bounded (because a grid of points can only represent a finite region). This limits application of the HMM approach to situations where the state space is naturally bounded.

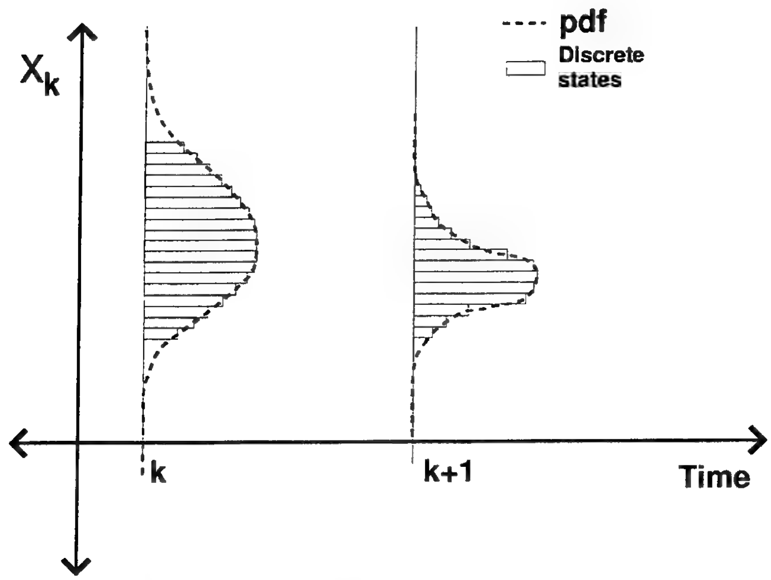

Figure 6 (U): Graphical interpretation of a HMM filter. The height of the bars represents the probability of being in the region of state space represented by the width of the bars.

Although there are no solid theoretical results to support the approximation of non-linear systems by HMMs there are several applications where this technique has been successfully applied. These applications include bearing only tracking problems [22], frequency tracking problems [23], and phase tracking problems.

7 Application: Target Tracking

The application presented in this section is motivated by work being done in the Guidance and Control group on the optimal precision guidance control problem. This control problem describes the terminal phase of a interceptor-target engagement. In this section we consider the filtering or state estimation problem related to the control problem.

For simplicity consider an engagement defined in continuous time and let the following

definitions be in a 2-D Euclidean frame. Let (x_t^I, y_t^I) and (x_t^T, y_t^T) be the position of the interceptor and target respectively. Then let (u_t^I, v_t^I), (u_t^T, v_t^T), (a_t^I, b_t^I) and (a_t^T, b_t^T) be the velocity and acceleration of the interceptor and target respectively.

Observations of the engagement are commonly related to the relative dynamics of the interceptor and target so we introduce the following state variable, $X_t := [x_t, y_t, u_t, v_t, a_t^T, b_t^T]$, where $x_t := x_t^T - x_t^I$ etc. The dynamics of the state can be expressed as follows

$$\begin{aligned}
\frac{dX_t}{dt} &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} X_t + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_t^I \\ b_t^I \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \cos\theta_t^T & -\sin\theta_t^T \\ \sin\theta_t^T & \cos\theta_t^T \end{bmatrix} \begin{bmatrix} \omega_t^{T,long} \\ \omega_t^{T,lat} \end{bmatrix} \\
\frac{dX_t}{dt} &= AX_t + Bu_t + G(X_t)\omega_t \qquad (7.1)
\end{aligned}$$

where $\theta_t^T = \tan^{-1}(v_t^T / u_t^T)$ is the target heading angle, $u_t := [a_t^I, b_t^I]'$, and $\omega_t := [\omega_t^{T,long}, \omega_t^{T,lat}]$. Although target acceleration is deterministically controlled by the target, in this model the target acceleration has been approximated by a "jinking" type model through the noises $\omega^{T,long}$ and $\omega^{T,lat}$. This acceleration model is simplistic (see [19] for more realistic target models) but is a reasonable representation in some situations.

Assume that the state is observed at evenly spaced distinct time instants $t_0, t_1, \ldots, t_k, \ldots$. Let index k denote the kth observation corresponding to the time instant $t = t_k$. Consider the following observation process

$$z_k = f(X_{t_k}, w_k^R, w_k^\theta) = \begin{bmatrix} R_k & + & R_k w_k^R \\ \theta_k^S & + & w_k^\theta \end{bmatrix} \qquad (7.2)$$

where $R_k = \sqrt{x_{t_k}^2 + y_{t_k}^2}$, $\theta_k^S = \tan^{-1}(y_{t_k}/x_{t_k})$ and w_k^R,w_k^θ are uncorrelated zero-mean Gaussian noises with variances σ_R^2 and σ_θ^2 respectively.

It is useful to consider a discrete-time representation of the continuous-time state equation (7.1) obtained through sampling theory. Let $h = t_k - t_{k-1}$ then using sample hold approximation the discrete-time state equation is

$$\begin{aligned}
X_{t_{k+1}} &= e^{Ah} X_{t_k} + G_{t_k}\bar{\omega}_k \quad \text{or} \\
X_{k+1} &= \bar{A} X_k + G_k \bar{\omega}_k \qquad (7.3)
\end{aligned}$$

where $\bar{\omega}_k = 1/h \int_{t_k}^{t_{k+1}} \omega_t dt$ and X_k etc. denote the discrete-time representation of X_k etc.. The variance of $\bar{\omega}_k$ is $1/h$ times the variance of ω_t.

We consider the engagement shown in Figure 7. The engagement commences at a distance of 5000 m. The interceptor is traveling at a velocity of 1000 m/s in a direction 36° (measured clockwise) from a line drawn between the interceptor and target. The interceptor is traveling at a velocity of 660 m/s in a direction of 120° from the same line. Hence, the initial conditions are $(x_0^I, y_0^I, u_0^I, v_0^I) = (0, 0, 1000\cos(36°), 1000\sin(45°))$ and $(x_0^T, y_0^T, u_0^T, v_0^T) = (5000, 0, 660\cos(120°), 660\sin(120°))$ where distances are in units of m and velocities are in units of m/s. Assume no control action is taken by the interceptor.

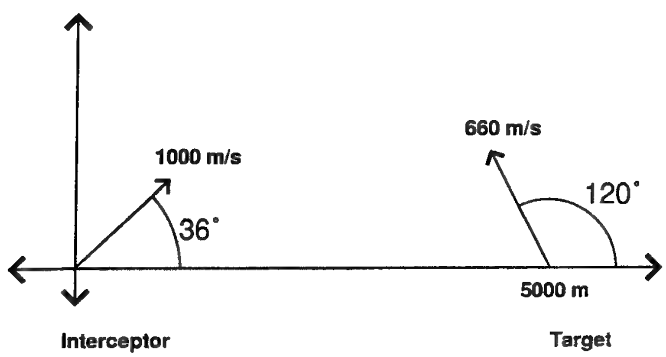

Figure 7 (U): Engagement configuration. The interceptor and target are roughly heading towards the same point (collision will not occur unless a manoeuvre is performed).

7.1 Extended Kalman Filter Approach

To apply the extended Kalman filter to this problem we obtain a linear approximation for the state equation. We approximate the non-linearity in the driving term as a time-varying linear functions (that is $A_k = \bar{A}$ and $B_k = G(\hat{X}_{k|k-1})$ where $\hat{X}_{k|k-1}$ is the one-step-ahead prediction of X_k). The measurement equation (7.2) is non-linear in the state and linearisation at $\hat{X}_{k|k-1}$ gives

$$\begin{aligned} C_k &= \left.\frac{\partial c_k(X)}{\partial X}\right|_{X=\hat{X}_{k|k-1}} \\ &= \begin{bmatrix} \hat{x}_{k|k-1}/\hat{R}_{k|k-1} & \hat{y}_{k|k-1}/\hat{R}_{k|k-1} \\ -\hat{y}_{k|k-1}/\hat{R}^2_{k|k-1} & \hat{x}_{k|k-1}/\hat{R}^2_{k|k-1} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{7.4}$$

The extended Kalman filter can now be implemented using the recursions (4.3) stated above.

The extended Kalman filter was used in a simulated engagement to estimate the target state at each time instant. The sampling rate of the simulation was 0.001 Hz. The initial estimate of the target position was $(5500, 500)$ m and velocity errors of 5 ms^{-1} in both x and y directions were present. The range measurement noise variance was $0.25R_k$, where R_k is the range at time k, and the angle measurement noise variance was 0.25. The target was assumed to be non-accelerating.

Figure 8 shows a plot of both the target and interceptor trajectories as well as the interceptor's estimate of the target's position. The initial position error quickly reduces and after 4.39 s, when the interceptor and target are 71 m apart (which is the closest distance achieved) the error in the estimated target position is 0.67 m. The extended Kalman filter also provides estimates of the target's velocity.

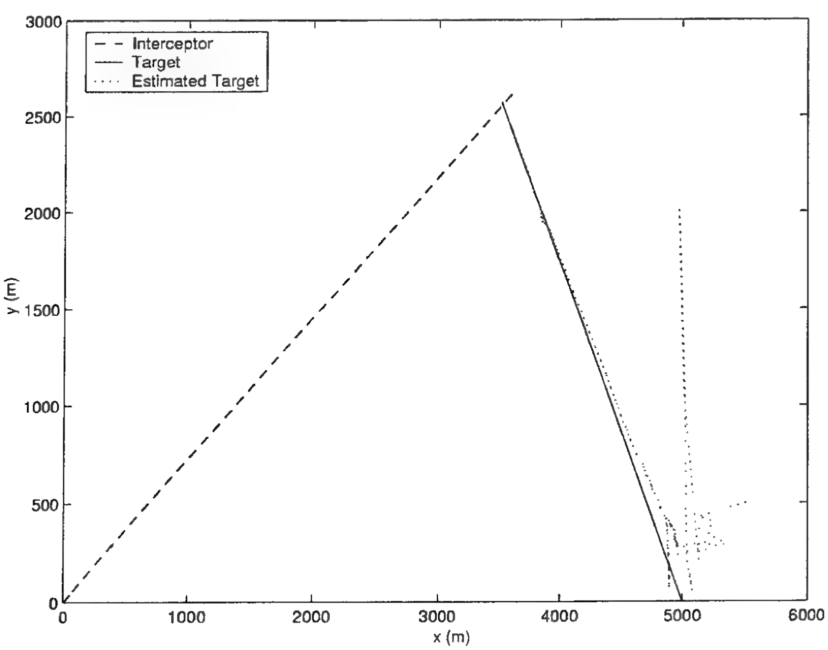

Figure 8 (U): Estimate Target Position

7.1.1 Stability of the Extended Kalman Filter

The stability of the extended Kalman filter in this situation can be examined using the results in Section 4.2. Because the state equations are linear the conditions for stability stated in Theorem 1 simplify to

$$\epsilon = \min\left(\epsilon_\chi, \frac{\underline{p}\underline{r}\alpha}{2\bar{p}^2\kappa_\chi}\left(2(1+\frac{\bar{p}}{\underline{r}})+\frac{\bar{p}\epsilon_\chi\kappa_\chi}{\underline{r}}\right)^{-1}\right) \tag{7.5}$$

where we have used $\bar{c}=1$, $\bar{a}=1$, ϵ_φ is unbounded and $\kappa_\varphi = 0$.

To determine whether stability of the extended Kalman filter can be assured for a particular initial error value we tested values of ϵ_χ in (7.5). We investigated stability against initial errors in the y position coordinate (assuming no error in x_0). Figure 9 shows the values of ϵ achieved for various values of ϵ_χ (note that this figure shows only the stability at the initial time instant and the stability of the filter at later time instants needs to be tested separately). From Figure 9, stability of the EKF can be guaranteed for initial errors in y less that 180 m. Using (7.5) it can be shown that when y_0 is known, errors in x_0 do not cause the EKF to diverge.

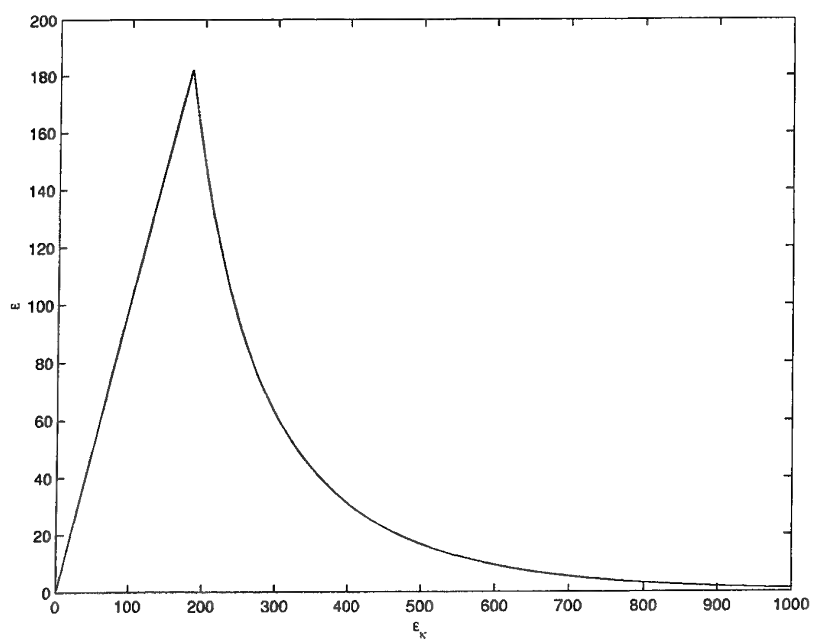

Figure 9 (U): The initial errors in y_0 for which the EKF is guaranteed to converge.

We know from the simulation results that the EKF converged from initial errors of 500m

in both axis. This demonstrates that although useful, the bounds produced by (7.5) are conservative.

7.2 Particle Filter and HMM Approaches

The particle filter was applied to the same target trajectory estimation problem. The noise variance was lowered to $0.05R_k$ and 0.05 in the range and angle measurements respectively. A system of 1000 particles was used to represent the evolution of the PDF. The initial support of the particle filter was sampled from a Gaussian density function with a mean displaced by the amount $(500, 500, 0.5, 0.5, 0, 0)$ from the true initial state with a covariance matrix $diag([500^2, 500^2, 0.5^2, 0.5^5, 10^{-6}, 10^{-6}])$. Here $diag(X)$ is the operation of creating a diagonal matrix from the vector X (if X is a vector) or the operation of making a vector from the diagonal of the matrix X (if X is a matrix).

A new support set was obtained at every 10 time instants by resampling from a Gaussian density function with the following mean and covariance:

$$\begin{aligned} \bar{x}_k &= \sum_{i=1}^{N^S} s_k^i w_k^i \\ var(\bar{x}_k) &= \sum_{i=1}^{N^S} (s_k^i - \bar{x}_k)(s_k^i - \bar{x}_k)' w_k, \end{aligned} \tag{7.6}$$

where $\bar{x}_k$ is an estimate of the state and $var(\bar{x}_k)$ is the co-variance matrix for $\bar{x}_k$.

The observation noise was inflated to included the effect state estimation errors in the following way:

$$\begin{aligned} var(R_k) &= 0.05R_k + C_1 var(\bar{x}_k)C_1' \\ var(\theta_k) &= 0.05 + C_2 var(\bar{x}_k)C_2' \end{aligned} \tag{7.7}$$

where C_1 and C_2 are the first and second rows of the linearisation matrix C used in the EKF.

The particle filter was able to estimate the trajectory when initialised on the true initial conditions; however, if initialised with any error, this error remained for the length of the simulation and the filter was unable to correct for this error. These simulation results suggests that when the initial PDF is not known the filter is neither divergent nor convergent but not very useful for this application.

HMM filter

To apply the HMM approach to this filtering problem requires discretisation of the state space. The HMM approach is not computationally tractable for the above target trajectory estimation problem. To illustrate the computational effort required consider the effort required for even the most basic and coarse approximation.

Assume that the velocity and acceleration are known (which is a fairly restrictive assumption in target tracking problems). For the configuration above, it is reasonable to bound the position space to $(-4500, 5500)$ in the x-axis and $(0, 1000)$ in the y-axis. If this space is coarsely discretised into 10 m by 10 m regions then 1×10^4 discrete states are required. The HMM filter on a 1×10^4 state process will require in the order of 1×10^8 calculations per time instant. Successful filtering is likely to require even finer discretisation than 10 m and hence computational requirements are likely to be significant less tractable.

8 Conclusions

This report presented a review of recent non-linear and robust filtering results for stochastic systems. Stability results for the extended Kalman filter and a robust Kalman filtering solution were presented. The report also examined a recent non-parametric filtering technique known as the particle filter. Finally, some simulation examples were presented that demonstrate the performance of several filters.

In many applications the extended Kalman filter may offer the best filtering solution, but in highly non-linear problems this filter is unlikely to perform well. For more non-linear problems there are various higher-order-model approaches that offer suboptimal filtering solutions. For the most complex problems, generic approaches such as the non-parametric particle filter may be appropriate, admittedly at a heavy computational cost.

All of these approaches assume certain knowledge of the state and measurement processes, which is unrealistic in practice. On the other hand, robust Kalman filtering and similar techniques mitigate for uncertainty in the system model, but this mitigation is generally at some performance loss.

Out of all these possible filtering approaches, no one approach is superior to the all others

and finding the most appropriate filter many require substantial investigation.

References

1. B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice Hall, New Jersey, 1979.

2. P.S. Maybeck, *Stochastic Models, Estimation, and Control*, Vol. 1, Academic Press, New York, 1979.

3. A.H. Sayed and T. Kailath, "A State-Space Approach to Adaptive RLS Filtering," *IEEE Signal Processing Magazine*, pp. 18-60, July 1994.

4. G.J. Bierman, "Measurement Updating Using the U-D Factorization," *Proc. of IEEE Conf. on Decision and Control 75*, Houston, Texas, pp. 337-346, 1975.

5. G.J. Bierman, "A comparsion of Discrete Linear Filtering Algorithms," *IEEE Trans. Aerospace and Electron. Systems*, Vol. 9, No. 1, pp. 28-37, 1973.

6. G.J. Bierman and C.L. Thornton, "Numerical Comparsion of Kalman Filter Algorithms: Orbit Determinination Case Study," *Automatica*, Vol. 13, pp. 23-35, 1977.

7. R.G. Brown and P.Y.C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 2nd Ed., John Wiley & Sons, New York, 1992.

8. A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic, New York, 1972.

9. B.F. La Scala, R.R. Bitmead and M.R. James, "Conditions for stability of the Extended Kalman Filter and Their Application to the Frequency tracking Problem," *Mathematics of Control, Signals, and Systems*, Vol. 8 pp. 1-26, 1995.

10. K. Reif, S. Günther, E. Yaz and R. Unbehauen, "Stochastic Stability of the Discrete-Time Extended Kalman Filter," *IEEE Trans. on Automatic Control*, Vol. 44, No. 4, April 1999.

11. S. Dey and J.B. Moore, "Risk-Sensitive Fitlering and Smoothing via Reference Probability Methods," *IEEE Trans. on Auotmatic Control*, Vol.42, No. 11, pp. 1587-1591, Nov. 1997.

12. J.J. Ford, "Risk-Sensitive Filtering and Parameter Estimation," Technical Report DSTO-TR-0764, DSTO ARML, Melbourne Australia, Jan. 1999.

13. R.K. Boel, M.R. James and I.R. Petersen, "Robust and Risk-Sensitive Filtering," *Proc. of Conference on Decision and Control*, CDC97, pp. 2273-2278, San Diego, California, Dec. 1997.

14. S. Dey and S. I. Marcus, "A framework for mixed estimation of hidden Markov models," *Proc. of Conference on Decision and Control*, CDC98, pp. 3473-3478, Tampa, Florida, Dec. 1998,

15. J. Carpenter, P. Clifford and P. Fearnhead, "Imporved Particle Filter for Nonlinear Problems," *IEE Proc. on Radar, Sonar and Navigation*, Vol. 146, No. 1, Feb. 1999.

16. A. Doucet, N. de Freitas, N. Gordon, "Sequential Monte Carlo Methods in Practice," Springer-Verlag, 2001.

17. D. Crisan, J. Gains and T. Lyons, "Convergence of a Branching Particle Method to the Solution of the Zakai Equation," *SIAM J. Appl. Math.*, Vol. 58, No. 5, pp. 1568-1590, Oct. 1998.

18. M.K. Pitt and N. Sheppard, "Filtering via Simulation: Auxiliary Particle Filters," to appear in *Journal of the American Statistical Association*, 1999.

19. Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*, Academic Press, Boston, 1988.

20. S.C. Krammer and H.W. Sorenson, "Recursive Bayesian estimation using piece-wise constant approximations," *Automatica*, Vol. 24, pp. 789-801, 1988.

21. R.J. Elliott, L. Aggoun, and J.B. Moore, *Hidden Markov Models: Estimation and Control*, Springer, New York, 1995.

22. A. Logothetis, R.J. Evans, L.J. Sciacca, "Bearings-Only Tracking Using Hidden Markov Models," *Proc. of CDC-94*, pp. 3301-3302, Lake Buena Vista, FL, Dec. 1994.

23. X. Xie and R.J. Evans, "Multiple Frequency Lne Tracking with Hidden Markov Models - Further Results," *IEEE Trans. on Signal Processing*, Vol. 41, No. 1, pp. 334-343, Jan. 1993.

24. I. R. Petersen and A. V. Savkin, *Robust Kalman filtering for signals and systems with large uncertainties*, Birkhuser, Boston Mass., 1999.

25. C.E. de Souza, M. Fu, and L. Xie, "H^∞ estimation for discrete-time linear uncertain systems," *International Journal or Robust and Nonlinear Control*, Vol. 1, No. 2, pp. 111-23, 1991.

26. J. Durbin and S. J. Koopman, "Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perpectives," *Journal of the Royal Statisitcal Society - Series B*, Vol. 62, Part. 1, pp. 3-56, 2000.

27. B.P. Carlin, N.G. Polson, D.S. Stoffer, "A Monte Carlo approach to nonnormal and nonlinear state space modeling," *Journal of the American Statisitcal Assocation*, Vol. 87, pp. 640-647, 1992.

28. H. Nijmeijer and A.J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer, New York, 1990.

29. I. Marceels and J. W. Polderman, *Adaptive Systems: An Introduction*, Birkhauser, Boston, 1996.

30. L. Ljung, *System Identification: Theory for the User*, Prentice Hall, London, 1987.

31. T.M. Cover and J.A. Thomas, *Elements of Infromation Theory*, John Wiley & Sons, New York, 1991.

DISTRIBUTION LIST

Non-linear and Robust Filtering: From the Kalman Filter to the Particle Filter

Jason J. Ford

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor

S&T Program

Chief Defence Scientist
FAS Science Policy
AS Science Corporate Management
Director General Science Policy Development

shared copy

Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Scientific Adviser Joint
Navy Scientific Adviser (Doc Data Sheet and distribution list only)

Scientific Adviser - Army (Doc Data Sheet and distribution list only)

Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory
Director

Chief of Weapons Systems Division
Research Leader Air Weapons Systems
Task Manager: Farhan Faruqi
Author: Jason Ford

DSTO Library and Archives

Library Fishermans Bend
Library Maribyrnong
Library Edinburgh 1 copy
Australian Archives
Library, MOD, Pyrmont
Library, MOD, HMAS Stirling
US Defense Technical Information Center, 2 copies
UK Defence Research Information Centre, 2 copies
Canada Defence Scientific Information Service, 1 copy

NZ Defence Information Centre, 1 copy

Capability Systems Staff

Director General Maritime Development (Doc Data Sheet only)
Director General Land Development
Director General Aerospace Development (Doc Data Sheet only)

Knowledge Staff

Director General Command, Control, Communications and Computers (DGC4) (Doc Data Sheet only)

Navy

SO (SCIENCE), COMAUSNAVSURFGRP, NSW (Doc Data Sheet and distribution list only)

Army

ABCA National Standardisation Officer, Puckapunyal (4 copies)
SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera QLD (Doc Data Sheet only)
NPOC QWG Engineer NBCD Combat Development Wing, Puckapunyal, VIC (Doc Data Sheet relating to NBCD matters only)

Air Force

Intelligence Program

DGSTA Defence Intelligence Organisation
Manager, Information Centre, Defence Intelligence Organisation

Corporate Support Program

Library Manager, DLS-Canberra
MS Sam Doran, Defence Library Service - Sydney West

Acquisitions Program

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Serials Section (M list), Deakin University Library, Geelong, VIC
Hargrave Library, Monash University (Doc Data Sheet only)
Librarian, Flinders University

OTHER ORGANISATIONS

NASA (Canberra)
AusInfo
State Library of South Australia

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

Library, Chemical Abstracts Reference Service
Engineering Societies Library, US
Materials Information, Cambridge Scientific Abstracts, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US
National Aerospace Laboratory, Japan
National Aerospace Laboratory, Netherlands

SPARES (5 copies)

Total number of copies:	**50**

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA						1. CAVEAT/PRIVACY MARKING
2. TITLE Non-linear and Robust Filtering: From the Kalman Filter to the Particle Filter			3. SECURITY CLASSIFICATION Document (U) Title (U) Abstract (U)			
4. AUTHOR Jason J. Ford			5. CORPORATE AUTHOR Aeronautical and Maritime Research Laboratory 506 Lorimer St, Fishermans Bend, Victoria, Australia 3207			
6a. DSTO NUMBER DSTO–TR–1301	6b. AR NUMBER AR 012-300		6c. TYPE OF REPORT Technical Report		7. DOCUMENT DATE April, 2002	
8. FILE NUMBER 9505-19-114	9. TASK NUMBER LRR 01/219	10. SPONSOR DSTO	11. No OF PAGES 39		12. No OF REFS 31	
13. URL OF ELECTRONIC VERSION http://www.dsto.defence.gov.au/corporate/reports/DSTO–TR–1301.pdf			14. RELEASE AUTHORITY Chief, Weapons Systems Division			

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DEFTEST DESCRIPTORS

Kalman filtering
Missile guidance
Signal analysis

19. ABSTRACT

This report presents a review of recent non-linear and robust filtering results for stochastic systems. We focus on stability and robustness issues that arise in the filtering of real systems. Issues such as numeric stability and the effect of non-linearity are also considered.

The report begins by introducing the famous Kalman filtering problem before proceeding to introduce the extended Kalman filter and related stability results. Robust forms of the Kalman filter and extended Kalman filter are also considered and finally a particle filtering approach is presented.

The report is intended to lead readers with a familiarity of the Kalman filtering problem through some of the more important recent (and not so recent) results on stability and robust filters in non-linear filtering problems.

Page classification: UNCLASSIFIED